

Inverse kinematics learning of a continuum manipulator using limited real time data

Alok Ranjan Sahoo, *Member, IEEE*, Pavan Chakraborty, *Member, IEEE*

Abstract—Data driven control of a continuum manipulator requires a lot of data for training but generating sufficient amount of real time data is not cost efficient. Random actuation of the manipulator can also be unsafe sometimes. Meta learning has been used successfully to adapt to a new environment. Hence, this paper tries to solve the above mentioned problem using meta learning. We consider two cases for that. First, this paper proposes a method to use simulation data for training the model using MAML(Model-Agnostic Meta-Learning). Then, it adapts to the real world using gradient steps. Secondly, if the simulation model is not available or difficult to formulate, then we propose a CGAN(Conditional Generative adversarial network)-MAML based method for it. The model is trained using a small amount of real time data and augmented data for different loading conditions. Then, adaptation is done in the real environment. It has been found out from the experiments that the relative positioning error for both the cases are below 3%. The proposed models are experimentally verified on a real continuum manipulator.

Index Terms—Continuum manipulator, Meta learning, Sim-to-real, Inverse kinematics

LACK of enough degrees of freedom and safety issues of the conventional manipulators lead researchers to the continuum manipulators. Snakes [1]–[3], tentacles [4]–[6], elephant trunks [7]–[10] were primary sources of inspiration for them. Mimicking their flexible designs is not an easy task.

Hence, researchers tried uniform backbone designs first. It was having simpler mathematical model but independent actuation of each section was difficult with it. Hence, tapered designs were introduced. Model based controllers were used for it. However, the mathematical formulations of it were complex. The properties of the manipulator must be known for it. Hence, data driven approaches were explored to reduce the difficulty.

In terms of model accuracy, it has been demonstrated that data-driven approaches perform better than model-based approaches. RNN based control [11], [12], reinforcement learning [13], deep reinforcement learning [14] are being used for the continuum manipulator. However, the requirement of high amount of data and time creates problem in using them in real world. These models are also not so effective to adapt to slight changes in the environment. To make our model truly intelligent, our model should be able to adapt to the new environment using a very small amount of changed data. Meta learning has shown significant success to generalize among similar types of tasks [15]. Hence, we will explore

its capability to generalize with a very small amount of real time data points.

It has been observed that meta learning based models perform well if it is trained with the data points which are similar to the real world scenario [16]–[18]. If we train it from scratch, then we will still need a considerable amount of real data points. Data points are generally generated using random actuation of the actuators. This can be unsafe and unpredictable some times. Generating real time data can also be costly. To address this problem generally the model is first trained with simulation data. Then, transfer learning [19], [20], knowledge distillation [21], imitation learning, meta learning are used for the its adaptation in real world. However, meta learning based sim-to-real adaptation has not been explored for continuum manipulators yet. Hence, we intend to explore meta learning for it in this paper. We will use MAML for it.

We segregated this problem in to two parts. First, if the simulation model is available, then we can use it for initial training. Then, we can adapt it using a meta learning(learning to learn) approach. Second, if the model is not available, then we can use a CGAN (Generative Adversarial Networks) model to create a data set from some small amount of real time data. Hence, we tried to find the effectiveness of the above two approaches for our highly non-linear continuum manipulator.

1) *A brief review of MAML:* Meta learning attempts to uncover generalisation after being trained with many sets of similar tasks, taking its cue from humans who tend to generalise after a few similar encounters. It is known as "learning to learn" since it tries to generalise across the questions. After training, it leverages this expertise (meta knowledge) to learn any task.

Meta learning can also be thought of as a bi-level optimization method. It alludes to a situation involving hierarchical optimization in which one optimization acts as a constraint on another optimization. The basic task such as classification or regression is done by the inner layer. The outer objective is taken care of by the meta learner. Task generalization performance or learning speed can be the outer objective function for the outer layer [22].

MAML(Model Agnostic Meta Learning) was proposed by finn et al. [15]. It was proposed for quick adaptation of a new task. It can be applied to any model trained with gradient descent. It can be applied to classification, regression, reinforcement learning etc. As we have already seen it becoming successful to adapt to a new condition in our previous work, we believe it has the potential to adapt to a real environment from simulation environment.

This paper was produced by the IEEE Publication Technology Group. They are in Piscataway, NJ.

Manuscript received April 19, 2021; revised August 16, 2021.

distribution over the tasks(τ). Let us consider that the model takes a single gradient step for the adaptation of the new task τ_i . Then the update equation can be represented as

$$\phi'_i = \phi - \alpha \nabla_{\phi} L_{\tau_i}(M_{\phi}). \quad (1)$$

Where, α = step size and L = loss function. The optimization is done for the model parameters based on the models performance with respect to different tasks sampled from $D(\tau)$.

$$\operatorname{argmin}_{\phi} \sum_{\tau_i \sim D(\tau)} L_{\tau_i}(M_{\phi'_i}) = \sum_{\tau_i \sim D(\tau)} L_{\tau_i}(M_{\phi - \alpha \nabla_{\phi} L_{\tau_i}(M_{\phi})}) \quad (2)$$

After that, model parameters are updated and the meta optimization is carried out using stochastic gradient descent.

$$\phi \leftarrow \phi - \beta \nabla_{\phi} \sum_{\tau_i \sim D(\tau)} L_{\tau_i}(M_{\phi'_i}) \quad (3)$$

Here, β is the meta step size. We will use Mean-squared error(MSE) as our loss function. Our L will be defined as

$$L_{\tau_i}(M_{\phi}) = \sum_{x^{(j)}, y^{(j)} \sim \tau_i} \|M_{\phi}(x^{(j)}) - y^{(j)}\|_2^2 \quad (4)$$

where $x^{(j)}$ and $y^{(j)}$ are an input/output pair sampled from task τ_i .

Continuum robot kinematics modelling requires two steps. First, a robot specific mapping between actuator space and configuration space is required. Then, the mapping between configuration space and task space is done. Given a target point, task space to actuator space mapping has to be done via configuration space. Here, we will try to find the mapping between task space to actuator space directly in the simulation environment and then, we will adapt it to the real world.

A. Problem formulation

We had used real data points for the initial training in our previous work [16]. However, it is not always possible to generate sufficient amount of real data points. Hence, we need to find a technique to reduce the requirement of real data. MAML has shown great success to adapt to new conditions and environments. Hence, we propose a MAML based technique to address the above stated problem. We will do the initial training with a simulation based model. Then, we will use small amount of real data points for adaptation of the real environment. Then, we propose a CGAN(Conditional Generative adversarial network) model for the data generation for the cases where developing the model of a manipulator is difficult.

To the best of the author's knowledge, it is the first work to use meta learning for the sim-to-real adaptation. Contributions of this paper are 1. A meta learning based sim-to-real adaptation technique for a continuum manipulator where the simulation model is available. 2. A CGAN(conditional generative adversarial network) model for generating data of different loading conditions. 3. A hybrid meta learning and generative adversarial network approach for training and adaptation of external loading was proposed.

Algorithm 1 Adaptation to the real environment

Step 1: Generate samples for learning in simulation environment using rand function $((p_{next}, a_{curr}, p_{curr}) \rightarrow a_{next})$
 Step 2: Train the model with simulation data using MAML
while not done **do**
 Generate samples for learning in real environment using rand function $((p_{next}, a_{curr}, p_{curr}) \rightarrow a_{next})$
 Update the model using the new K data samples
end while

As explained in our previous work [16], a position level controller can be formulated using the following equation

$$J(a_i)a_{i+1} \approx p_{i+1} - \psi(a_i) + J(a_i)a_i \quad (5)$$

The ψ function in this case gives a mapping from the actuator space to the task space. The task space is $p \in R^n$ and the actuator space is $a \in R^m$. a_{i+1} and a_i represent the present and next actuator positions respectively. Current and next position of the end effector are represented by p_i and p_{i+1} . As feed-back to a controller has shown higher success then open loop, we will learn a mapping between $(p_{i+1}, a_i, p_i) \rightarrow a_{i+1}$. Similar to our previous work, our model will explore the whole task space. Whole range of actuators will be used for data collection.

I. ARCHITECTURE OF THE NEURAL NETWORK

There are three hidden layers of size 128 in our neural network. There are 11 nodes in the input layer. Target tip position, present tip position, present actuator position, and external load are the inputs. The output layer is having 4 nodes. The output is the target actuator positions. ReLU was used as the activation function. The optimizer used for it was Adam. Both inner loop step size and meta step size are kept at 0.01. The loss function has defined as:

$$L(M) = \sum_{x^{(j)}, y^{(j)} \sim \tau_i} \|(M(x^{(j)}) - y^{(j)})\|_2^2 \quad (6)$$

Here, $x^{(j)}$ is the tensor of (p_{i+1}, a_i, p_i) , load and $y^{(j)}$ is the a_{i+1} (target tip positions). The loss function is the MSE between the predicted actuator locations and the actual actuator positions needed to reach a specific target.

II. SIMULATION ENVIRONMENT

Here, we have used 3d mechanics based model by Camarillo et al. [23] for the simulation. It was used because this is capable to accommodate variable length backbone. As our prototype is a spring based continuum manipulator, the backbone length varies with respect to the actuation. In our previous work, we had shown the procedure to adapt the model for our prototype. We have used the parameters obtained from the characterization of our manipulator in our previous work. Here, to generate the data points, we have used the forward kinematics of the above mentioned model. The external loading has been incorporated using the tension in the string as:

$$T_{externalloading} = T_{zeroloading} + externalload/4$$

TABLE I
HYPER PARAMETERS USED

Step size	0.01
Meta step size	0.01
k	10
Meta batch size	20
Number of hidden layers	3
Hidden layer size	128
Epochs	200

A. Simulation data collection

We collected 10,000 different input/output pairs for external loading of 0 to 1kg(with a resolution of 0.1 kg). Rand function was used for random actuation of the actuators. Initially, we maintained a restriction that difference between the actuator position of a particular actuator in consecutive actuation should be less than 10% of the maximum actuator length(0.25 m). We changed the actuator values to corresponding encoder values to match the input types of the real robot. After 10% data collection, we removed this restriction. The data set was segregated into training, validation set and test set in the ratio of 70:15:15.

III. EXPERIMENTAL SETUP

We have used our spring based continuum manipulator for our experimentation [16], [24]. The manipulator is 0.77 long. It is having 4 sections. However, only section 2 and 4 are active sections here. The spring back-boned manipulator has a pvc skin. our previous work can be referred for the complete description of the experimental set up and the manipulator [16], [24] . We will be using the distal section for our experimentation. Table IX shows the parameters used for the training.

IV. COLLECTING DATA FROM THE PROTOTYPE

We discretized the actuator space with respect to the encoder values to collect the data points from the prototype. Here, we gave 101 random actuator values for each external loading in the range of [0,0.5] with a difference of 0.05 kg. The actuator values and corresponding tip point values were recorded for the trajectories. 20 values in regular interval between the starting input-output pairs were treated as the data points for training(excluding starting point). Hence, we have 2000 data points for each external loading condition. In the following experiments, we will see the number of gradients required for a particular external loading to adapt to that external load in the real world environment. Here, we will use an external load of 0.2 kg on the prototype.

V. EXPERIMENTS AND RESULTS

A. Testing with the single section model

To test the capability of adaptation, we first trained the model with a single section continuum manipulator. The simulation model used the parameters of the distal section of the prototype. The length of the manipulator was taken to be 0.77 m. The model was a single section uniform cross

section and uniform stiffness continuum manipulator. 10,000 data points were collected for the training as described in the previous section. We took random 30 points in the real world to find the average error. For adaptation purpose, we will do gradient update after 50 samples of input-output pair of the real environment. To know the actual achievable accuracy, we tested a model trained with the real data points. The accuracy achieved with the real data point was 0.0210 ± 0.0119 m. We will use this as the default accuracy.

TABLE II
RESULTS OF THE MODEL TRAINED WITH REAL DATA

No. of data points	Average error(in m)	standard deviation(in m)
0	0.0210	0.0111

TABLE III
RESULTS OF THE MODEL TRAINED WITH UNIFORM CONTINUUM MANIPULATOR DATA(SINGLE SECTION)

No. of gradient steps	Average error(in m)	standard deviation(in m)
0	0.0839	0.0254
1	0.0625	0.0233
2	0.0543	0.0192
3	0.0522	0.0213
4	0.0495	0.0189
5	0.0416	0.0173
6	0.0358	0.0155
7	0.0359	0.0154

Here, from the table III, we can observe that model trained with the single section data can achieve a positioning error up to 4.64%. We had seen this type of initial error in our last work. There, we had put an external load above the allowed external load (0.5 kg)(table ??). The model could reduce the error up to 6.8% in that case. However, in this case it could reduce the error up to 4.64%. Hence, we tried to reduce the error further by using a double section model.

B. Testing with the double section model

Here, the length of each section was 0.38m. The distal section inherited the parameters of the distal section of the prototype. Similarly, the proximal section of the simulation model inherited the parameters of the second section. We again followed the same procedure as followed for the single section case.

TABLE IV
RESULTS OF THE MODEL TRAINED WITH TWO SECTION CONTINUUM MANIPULATOR DATA

No. of gradient steps	Average error(in m)	standard deviation(in m)
0	0.0774	0.0244
1	0.0406	0.0248
2	0.0344	0.0181
3	0.0289	0.0139
4	0.0256	0.0128
5	0.0209	0.0115

Table IV describes the results obtained for different number of gradient steps. We can observe from the table that the initial

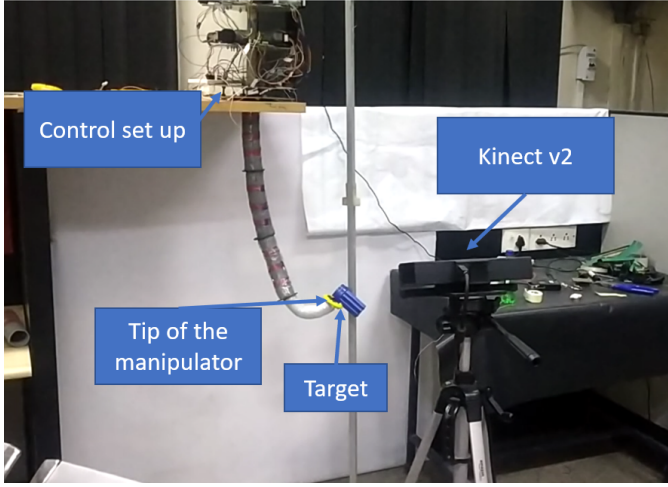


Fig. 1. Target reaching by the manipulator in real environment

error of the model is slightly less than the previous model. However, it could reduce the relative mean error to the range of 2.75% with 5 gradient steps. Hence, we can say that we can use the controller trained with the double section model in real world .

C. Random point target with different loading condition

Now, let us see the capability of the model to adapt to the external load in the real world directly after the adaptation step in the last experiment. We will give 50 random points as the targets for each external loading. Then, we will find the average error for it.

TABLE V
RESULTS FOR THE DIFFERENT LOADING CONDITIONS

External loading (in kg)	Average error (in m)	standard deviation (in m)
0	0.0211	0.0139
0.1	0.0214	0.0124
0.2	0.0207	0.0113
0.3	0.0221	0.0133
0.4	0.0216	0.0141
0.5	0.0219	0.0138

From the table V we find that the model performs equally well for all the loading conditions. This gave best result for 0.2 kg case as the model had adapted to the same external loading. However, the average error of the model did not have much difference between different loading conditions. It shows that if a model is trained with data points of different loading conditions, then it will perform much better than the approach we took in our last work. Figure 1 shows the capability of the manipulator to reach the target.

D. Trajectory following experiment with known loading

To see the capability of the model trained with the simulation data, we will consider two different trajectories. The first one will be a straight line and second one will be a semi circle. The length of the straight line is 0.60 m. We took 18 control points in between starting and ending positions with

TABLE VI
RESULTS OF TRAJECTORY FOLLOWING TASKS WITH KNOWN EXTERNAL LOADING

Task No.	Loading(in kg)	Tip positioning error(in m)	Standard deviation(in m)
Trajectory 1	0.2	0.0218	0.0097
Trajectory 2	0.2	0.0227	0.0106

TABLE VII
TRAJECTORY FOLLOWING EXPERIMENT WITH UNKNOWN EXTERNAL LOADING

Task No.	Loading(in kg)	Tip error(in m)	Standard deviation(in m)
Trajectory 1	0.250	0.0231	0.0102
Trajectory 2	0.250	0.0237	0.0098

an interval of 0.03 m. The radius of the semicircle was chosen to be 0.3m. Control points at an interval of 10 degrees were given for it. Here, we have used an external loading of 0.2 kg.

From the figure 2, we observe that the model took slightly zig-zag line. This phenomenon can be attributed to the fact that the distance between the control points are slightly larger. However, the average error for it remained below 0.0220 m. In the second trajectory, we observe the error was more than the straight line case. However, it was below 0.0230 m(relative error of below 3%). So, we can see that our model is performing well for known loading case.

E. Trajectory following experiment with unknown external loading

To test the capability of the model for unknown loading case, we used an external load of 0.250 kg. We used the same trajectory for this case as well.

From the figure 4 and 5, it can be observed that we see a similar pattern as in case of known loading case. Here, the error was higher than the known case. However, the relative error is still below 3%.

VI. A COMPARATIVE STUDY

We compared our result with the BPNN model proposed by Thuruthel et. al [25] as we did in the last work. We trained the model with 10000 training data as before. Then , we tested it for the unknown loading case.

From the figure 6, it is evident that the model is performing poorly in the real environment. The error shoots above 0.06

TABLE VIII
COMPARATIVE RESULTS FOR THE TRAJECTORY 2 WITH UNKNOWN EXTERNAL LOAD

Algorithm	Loading(in kg)	Tip error(in m)	Standard deviation(in m)
BPNN	0.25	0.0491	0.0161
MAML based adaptation	0.25	0.0237	0.0098
MAML based adaptation from CGAN data	0.25	0.0256	0.0102

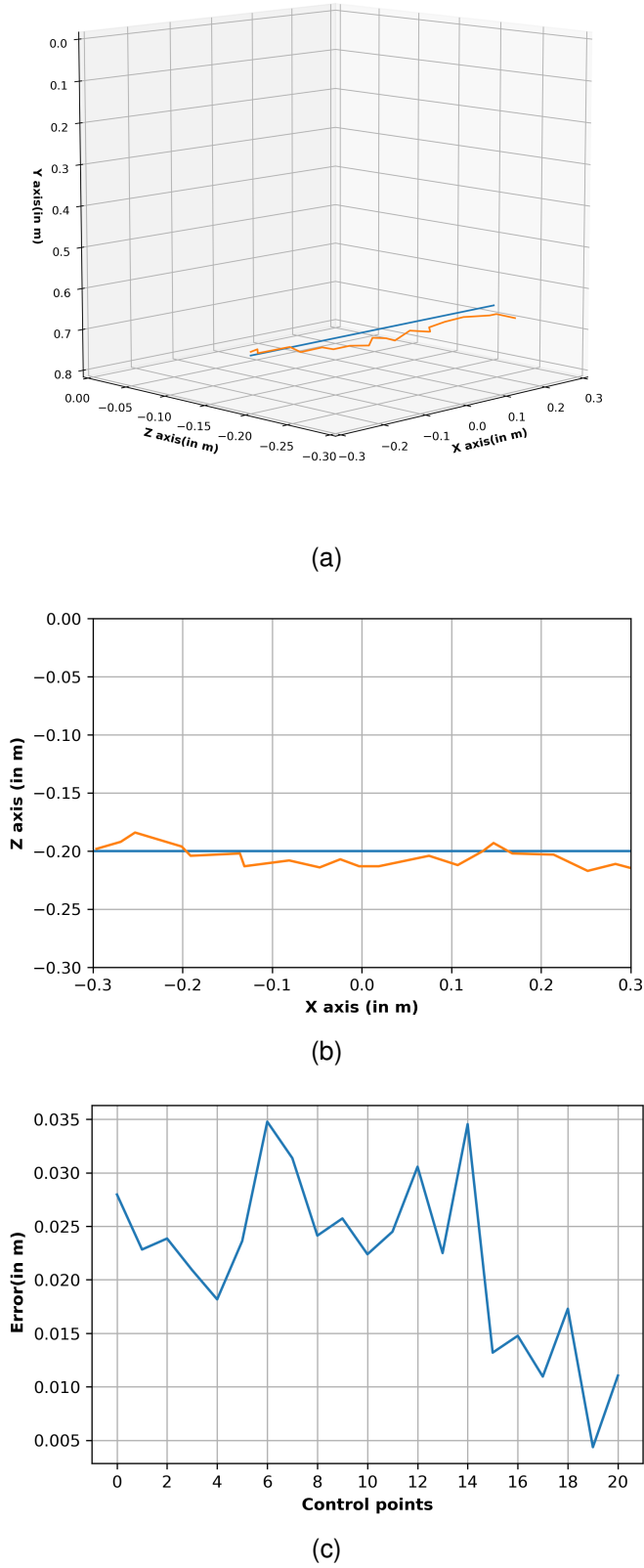


Fig. 2. Graph of trajectory 1 with known loading (a) 3D view (b) View in ZX plane (c) Tip positioning error. Blue and orange legends show desired trajectory and achieved trajectory respectively

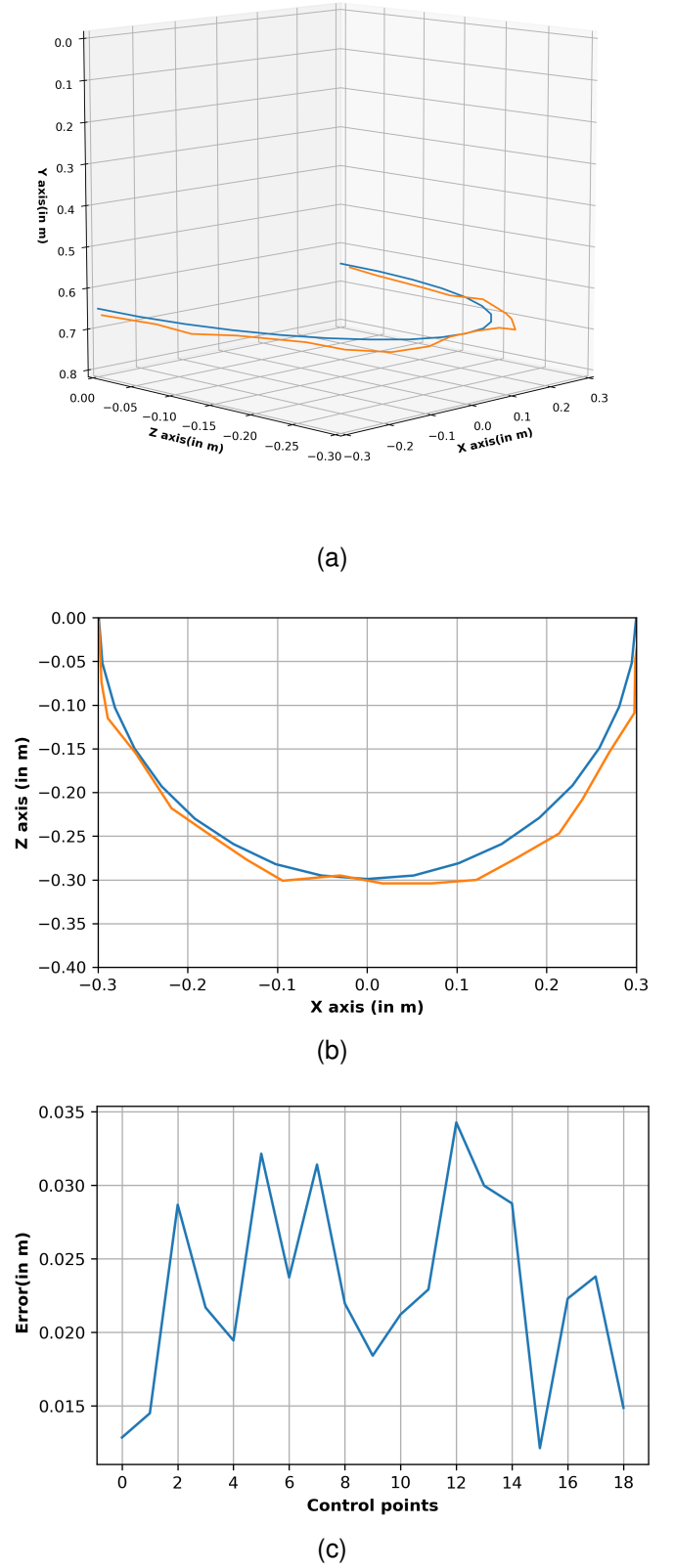
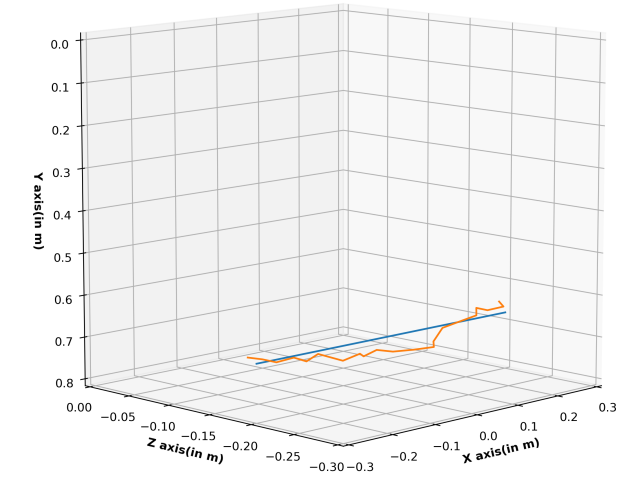
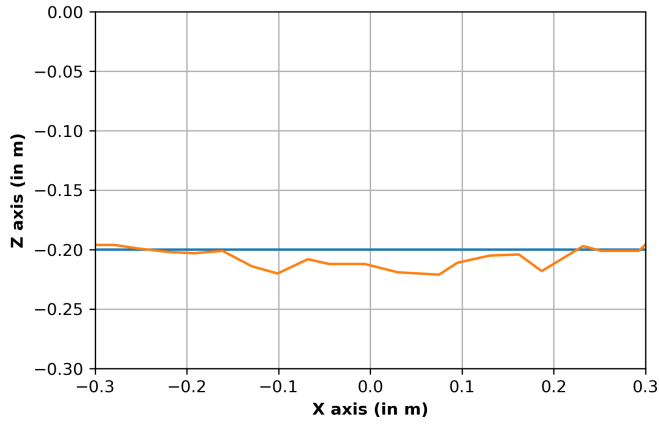


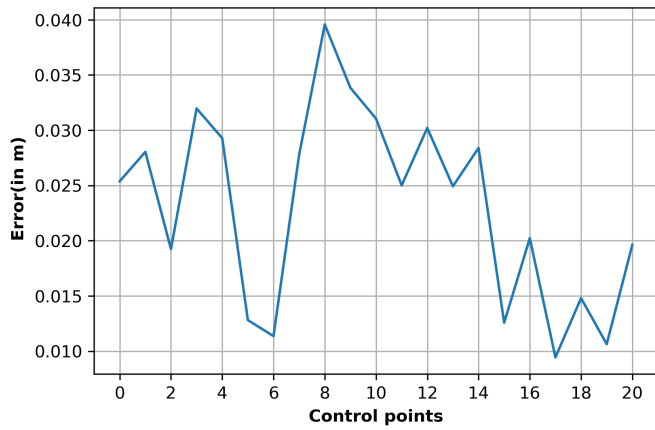
Fig. 3. Graph of trajectory 2 with known loading (a) 3D view (b) View in ZX plane (c) Tip positioning error. Blue and orange legends show desired trajectory and achieved trajectory respectively



(a)

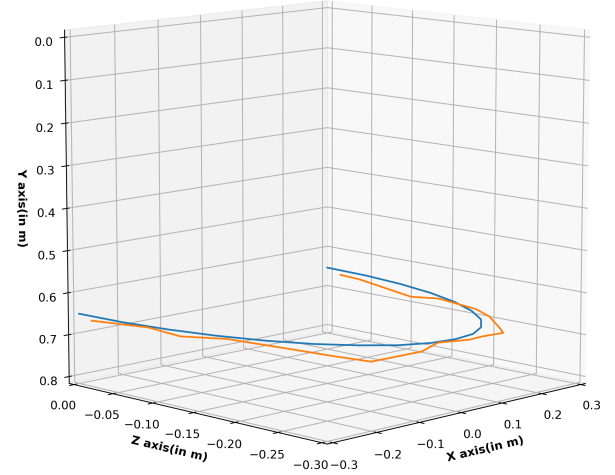


(b)

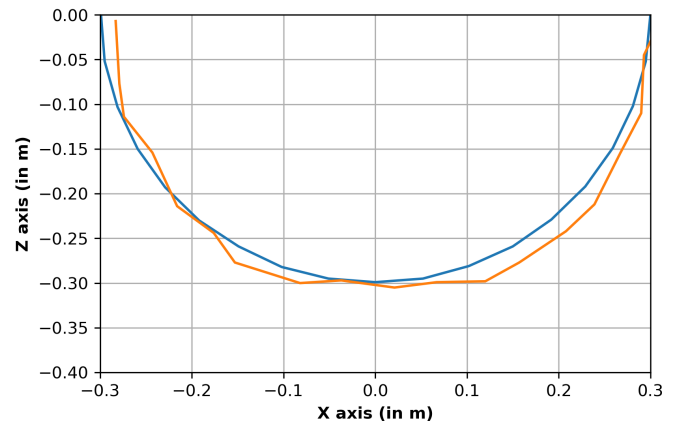


(c)

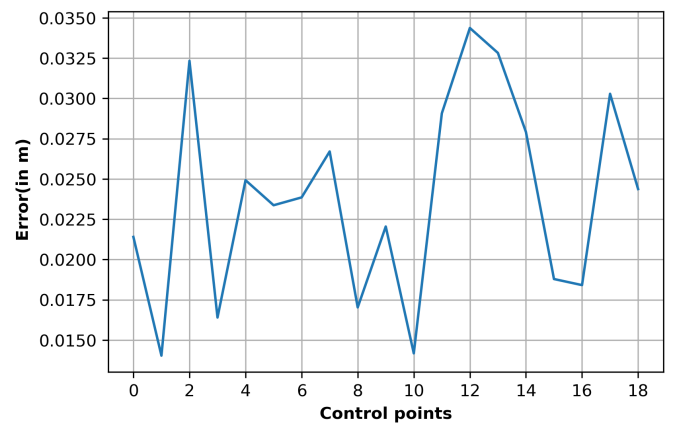
Fig. 4. Graph of trajectory 1 with unknown loading (a) 3D view (b) View in ZX plane (c) Tip positioning error. Blue and orange legends show desired trajectory and achieved trajectory respectively



(a)



(b)



(c)

Fig. 5. Graph of trajectory 2 with unknown loading (a) 3D view (b) View in ZX plane (c) Tip positioning error. Blue and orange legends show desired trajectory and achieved trajectory respectively

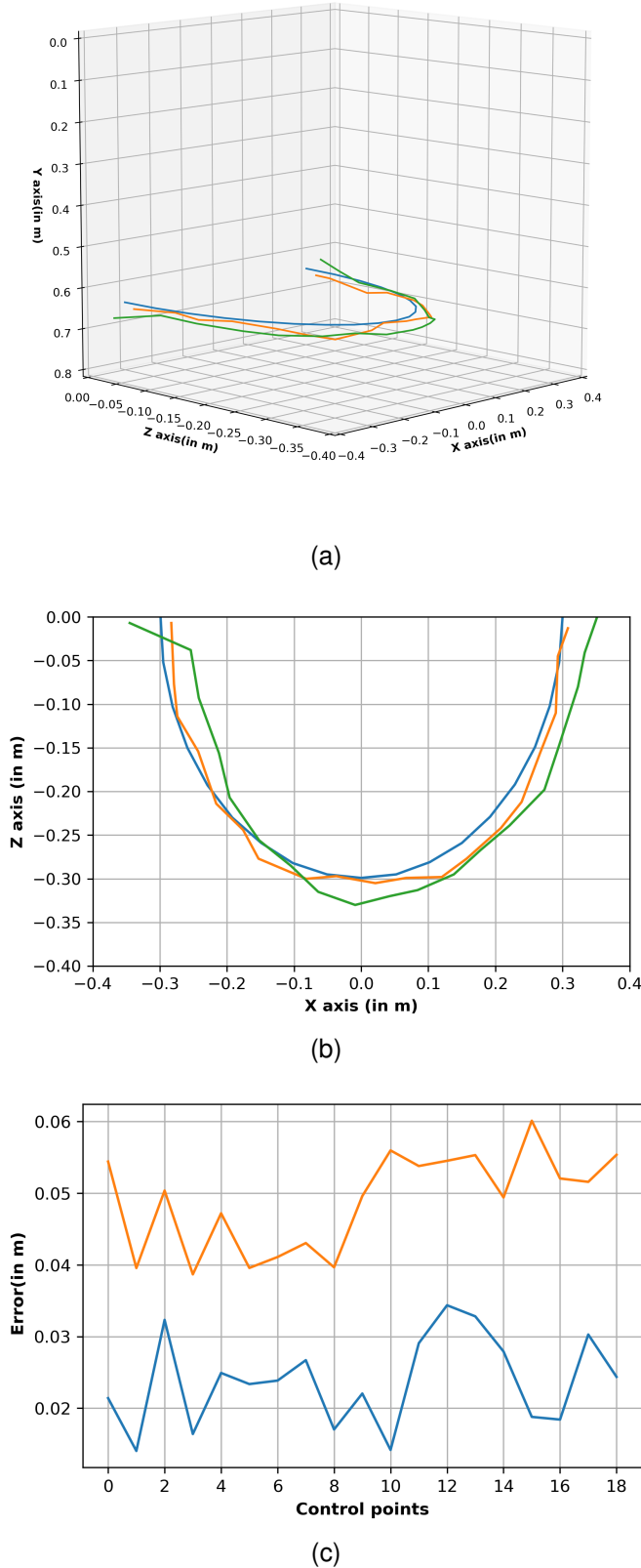


Fig. 6. Graph of trajectory 2 with unknown loading (a) 3D view (b) View in ZX plane (c) Tip positioning error. Blue, orange and green legends show desired trajectory, achieved trajectory and achieved trajectory by our controller and BPNN respectively.

m. The average error almost touches 0.05m. It could not adapt to the unknown load unlike last time where it tried to adapt to the new loading conditions. Now, it can be safely said that our model outperforms the BPNN model by a huge margin.

VII. USING GAN FOR DATA GENERATION

From the above experiments, we saw that the simulation data is quite helpful for the initial training of the model. If we don't have a simulation model, then getting real data becomes the only option. GAN has currently shown a lot of success in creating fake datas in computer vision. Hence, we propose that a GAN can be used to create similar data sets using small number of real data sets. We are aware of the fact that GAN requires a huge amount of data to give good accuracy. However, we will try to train the model using small number of data as getting higher accuracy is not the main target here. The idea here is that we will use it for initial training only. We will use the same adaptation technique as discussed earlier.

A. A brief review of GAN

Generative Adversarial Networks(GAN)s have shown great success in computer vision implementations. However, limited implementations have been done in continuum manipulator control. The basic idea of GAN is to let two neural networks (generator and discriminator) compete against each other. The work of the generator is to create fake data similar to the original data. The job of the discriminator is to differentiate between real and fake data.

Let us consider the original data distribution to be p_o . The generator should be able to find a data distribution p_g close to the p_o . It is generally hard to get p_r directly. Hence, this approach helps to get a proper data distribution [26].

Let us consider the discriminator and the generator have parameters Φ_d and Φ_g . First, the discriminator $Disc(x; \Phi_d)$ is trained using labeled real data and fake data. After it is able to differentiate, the generator, $Gen(x; \Phi_g)$, takes input from the latent space z . It tries to map it to a sample in the distribution of p_g . Then, the discriminator tries to differentiate between the real one and fake one. Based on the results from the discriminator, the generator tries to create more realistic fake data by giving a probability value $p(x; \Phi_g)$. The over all objective function for it is [27]

$$\min_{Gen} \max_{Disc} V(Disc, Gen) = E_{x \sim p_{data}(x)} [\log p(x)] + E_{x \sim p_{model}} [\log(1 - p(x))] \quad (7)$$

Here, $V(Disc, Gen)$ determines the payoff of the discriminator. Generator receives $-V(Disc, Gen)$ as its own payoff. This converges when the discriminator is no longer capable to differentiate between real and fake data. It outputs 0.5 then. After convergence, generator is used to create fake data.

However, it is seen that GANs are quite unstable and it may suffer from vanishing gradient problem. It also collapses where the generator outputs nearly similar output.

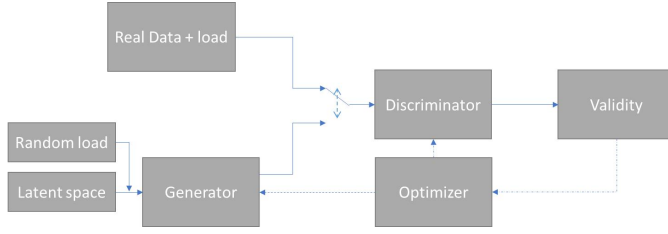


Fig. 7. Proposed Conditional GAN

TABLE IX
DETAILS OF GAN NETWORK TRAINING

No. of layers	2
Hidden layer size	256
Activation function	ReLU
Epochs	200
loss	binary cross-entropy
optimizer	adam
Latent space dimension	50

B. CGAN

The main idea of conditional GAN is that a support information vector y is provided here. It is available to both discriminator and generator. It helps to converge faster. Moreover, we can have the required output by the generator by using the information. In our case, the information vector is the external load. Here, we first gave 13 random actuator values for each external loading in the range of $[0,0.5]$ with a difference of 0.1 kg. The actuator values and corresponding tip point values were recorded for those trajectories. 30 values in regular interval between the starting input-output pairs were treated as the data points for training(excluding starting point). Given those points, we used a simple technique to increase the data points. We created random values between $[-50000,50000]$. We added those randomly to the existing encoder values. We also changed the tip position by randomly adding values between $[-0.01,0.01]$ randomly in x, y and z directions. We increased the number of data points 6 times using this technique.

C. Error in the CGAN data

To test the accuracy of the generated data points, we took 50 data points having $((p_{next}, a_{curr}, p_{curr}) \rightarrow a_{next})$ values. To verify its accuracy, we did a random point test. We took 50 random data points generated by the generator. Then, we tested the tip positioning error for it on the real robot. We found the average error to be $0.0351 \pm 0.0115m$. We generated 8000 data points from the generator. Now, we used these data set with the original data set to train the neural network.

D. Experiment with CGAN-MAML model

We used the controller directly to the real environment first. It was found out that the average error for it was 0.0336 ± 0.0193 at first. It is comparatively less than the result we got from the model trained with simulation data. However, from the table X, we find that the current model is slower in adapting to the actual world in comparison to the previous

TABLE X
RESULTS OF THE THE MODEL TRAINED WITH CGAN DATA

Number of gradient steps	Average error(in m)	standard deviation(in m)
0	0.0336	0.0198
1	0.0289	0.0172
2	0.0247	0.0185
3	0.0235	0.0135
4	0.0232	0.0129

model trained with simulated data. It may be due to the fact that the previous model was trained with the data having higher distribution. The current model was trained using more similar data. Hence, it is always better to have a simulation model. However, if the simulation model is not available and collecting real time data is expensive, then we can use the above proposed method.

E. A comparative study

We further tried to see its performance with respect to previous controllers. Hence, we choose to see its performance for trajectory following task. We choose trajectory-2 for it. Here, we will consider unknown loading case. From the figure 8, we can observe that model trained with data from cgan performed better than BPNN model from the last experiment. However, it could not surpass the performance of the model trained with simulation data. As explained earlier, it must be due to the models training using similar data. However, the average error was below 0.03m (0.0256 m). Hence, it can be used for the cases where we do not have a simulation model. We can further improve the model's accuracy using more data points to train the CGAN. Further experiments can be done to find out exact ratio of real to synthetic data needed to achieve better accuracy.

VIII. DISCUSSIONS AND CONCLUSIONS

In this paper, we proposed a MAML based sim-to-real approach. First, we trained the controller using simulation model based on 3D mechanics model of the robot. Then, we used the adaptation step to adapt to the real world. It was observed that the model trained with only simulation data points required 5 gradient steps to adapt to the real environment. The relative errors were below 2.94% for known cases and 3.07 % for unknown loading cases. Thuruthel et. al [11], [25] were successful to achieve relative error up to 2.76 percent using model based reinforcement learning [28]. Deep reinforcement learning based controller of Satheeshbabu et. al [29] achieved a relative error of 3.87 percent. We further compared the model's performance with BPNN model. It was found that our model out performs the BPNN model. Then, we proposed a CGAN based method for the cases where simulation model is not available. We trained the CGAN model with the data from generated 1800 real data points. Then, we generated fake data for the training using CGAN. After adaptation step, the relative error for it was below 3%. Finally, we can come to the conclusion that we can use both simulator data and cgan generated low accuracy data for training and use MAML based adaptation in the real world.

REFERENCES

- [1] D. Shen, Q. Zhang, C. Wang, X. Wang, and M. Tian, "Design and analysis of a snake-inspired crawling robot driven by alterable angle scales," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3744–3751, 2021.
- [2] I. Virgala, M. Kelemen, E. Prada, M. Sukop, T. Kot, Z. Bobovský, M. Varga, and P. Ferenčík, "A snake robot for locomotion in a pipe using trapezium-like travelling wave," *Mechanism and Machine Theory*, vol. 158, p. 104221, 2021.
- [3] K. Y. Pettersen, "Snake robots," *Annual Reviews in Control*, vol. 44, pp. 19–44, 2017.
- [4] J. Jørgensen, K. B. Bojesen, and E. Jochum, "Is a soft robot more "natural"? exploring the perception of soft robotics in human–robot interaction," *International Journal of Social Robotics*, vol. 14, pp. 95–113, Jan 2022.
- [5] J. Fras, M. Macias, Y. Noh, and K. Althoefer, "Fluidical bending actuator designed for soft octopus robot tentacle," in *2018 IEEE International Conference on Soft Robotics (RoboSoft)*, pp. 253–257, IEEE, 2018.
- [6] R. V. Martinez, J. L. Branch, C. R. Fish, L. Jin, R. F. Shepherd, R. M. Nunes, Z. Suo, and G. M. Whitesides, "Robotic tentacles with three-dimensional mobility based on flexible elastomers," *Advanced materials*, vol. 25, no. 2, pp. 205–212, 2013.
- [7] M. W. Hannan and I. D. Walker, "Kinematics and the implementation of an elephant's trunk manipulator and other continuum style robots," *Journal of robotic systems*, vol. 20, no. 2, pp. 45–63, 2003.
- [8] K. Wu and G. Zheng, "Fem-based gain-scheduling control of a soft trunk robot," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3081–3088, 2021.
- [9] T. Mahl, A. Hildebrandt, and O. Sawodny, "A variable curvature continuum kinematics for kinematic control of the bionic handling assistant," *IEEE Transactions on Robotics*, vol. 30, no. 4, pp. 935–949, 2014.
- [10] J. Yang, E. P. Pitarch, J. Potratz, S. Beck, and K. Abdel-Malek, "Synthesis and analysis of a flexible elephant trunk robot," *Advanced Robotics*, vol. 20, no. 6, pp. 631–659, 2006.
- [11] T. G. Thuruthel, E. Falotico, M. Manti, and C. Laschi, "Stable open loop control of soft robotic manipulators," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 1292–1298, 2018.
- [12] T. G. Thuruthel, E. Falotico, F. Renda, and C. Laschi, "Learning dynamic models for open loop predictive control of soft robotic manipulators," *Bioinspiration & biomimetics*, vol. 12, no. 6, p. 066003, 2017.
- [13] X. You, Y. Zhang, X. Chen, X. Liu, Z. Wang, H. Jiang, and X. Chen, "Model-free control for soft manipulators based on reinforcement learning," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2909–2915, IEEE, 2017.
- [14] S. Satheeshbabu, N. K. Uppalapati, G. Chowdhary, and G. Krishnan, "Open loop position control of soft continuum arm using deep reinforcement learning," in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 5133–5139, IEEE, 2019.
- [15] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *Proceedings of the 34th International Conference on Machine Learning (D. Precup and Y. W. Teh, eds.)*, vol. 70 of *Proceedings of Machine Learning Research*, pp. 1126–1135, PMLR, 06–11 Aug 2017.
- [16] A. R. Sahoo and P. Chakraborty, "A study on position control of a continuum arm using maml (model-agnostic meta-learning) for adapting different loading conditions," *IEEE Access*, 2022.
- [17] W. Zhao, J. P. Queralta, and T. Westerlund, "Sim-to-real transfer in deep reinforcement learning for robotics: a survey," in *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 737–744, IEEE, 2020.
- [18] K. Arndt, M. Hazara, A. Ghadirzadeh, and V. Kyriki, "Meta reinforcement learning for sim-to-real domain adaptation," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2725–2731, 2020.
- [19] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2020.
- [20] K. Tanaka, R. Yonetani, M. Hamaya, R. Lee, F. von Drigalski, and Y. Ijiri, "Trans-am: Transfer learning by aggregating dynamics models for soft robotic assembly," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4627–4633, IEEE, 2021.
- [21] A. A. Rusu, S. G. Colmenarejo, C. Gulcehre, G. Desjardins, J. Kirkpatrick, R. Pascanu, V. Mnih, K. Kavukcuoglu, and R. Hadsell, "Policy distillation," *arXiv preprint arXiv:1511.06295*, 2015.
- [22] T. M. Hospedales, A. Antoniou, P. Micaelli, and A. J. Storkey, "Meta-learning in neural networks: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2021.

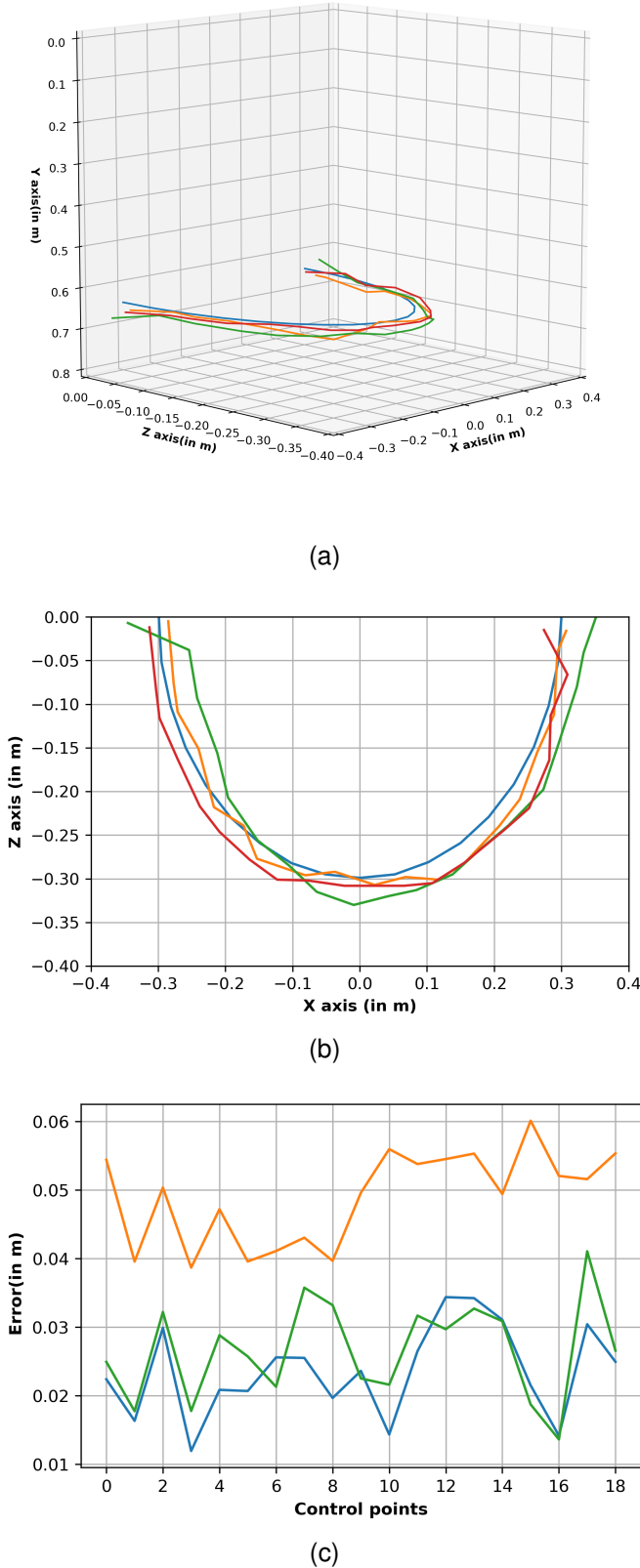


Fig. 8. Graph of trajectory 2 with unknown loading (a) 3D view (b) View in ZX plane (c) Tip positioning error . The blue, orange, green and red legends show the desired trajectory, achieved trajectory using our controller, BPNN and cgan data trained controller respectively.

- [23] D. B. Camarillo, C. R. Carlson, and J. K. Salisbury, "Configuration tracking for continuum manipulators with coupled tendon drive," *IEEE transactions on robotics*, vol. 25, no. 4, pp. 798–808, 2009.
- [24] A. R. Sahoo and P. Chakraborty, "Development and analysis of a bio-inspired wire-driven variable stiffness double spring based tapered multi-section flexible robot," *Industrial Robot: the international journal of robotics research and application*, 2021.
- [25] T. George Thuruthel, E. Falotico, M. Manti, A. Pratesi, M. Cianchetti, and C. Laschi, "Learning closed loop kinematic controllers for continuum manipulators in unstructured environments," *Soft robotics*, vol. 4, no. 3, pp. 285–296, 2017.
- [26] N. Gao, H. Xue, W. Shao, S. Zhao, K. K. Qin, A. Prabowo, M. S. Rahaman, and F. D. Salim, "Generative adversarial networks for spatio-temporal data: A survey," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 13, no. 2, pp. 1–25, 2022.
- [27] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.
- [28] T. G. Thuruthel, E. Falotico, F. Renda, and C. Laschi, "Model-based reinforcement learning for closed-loop dynamic control of soft robotic manipulators," *IEEE Transactions on Robotics*, vol. 35, no. 1, pp. 124–134, 2018.
- [29] S. Satheshebabu, N. K. Uppalapati, T. Fu, and G. Krishnan, "Continuous control of a soft continuum arm using deep reinforcement learning," in *2020 3rd IEEE International Conference on Soft Robotics (RoboSoft)*, pp. 497–503, IEEE, 2020.



ALOK RANJAN SAHOO received the B.Tech degree in Mechanical engineering. He received the M.Tech degree in Robotics from IIIT Allahabad. He is currently pursuing PhD in Robotics from IIIT Allahabad. His research interests are robotics, machine learning and computer vision.



PAVAN CHAKRABORTY received the Ph.D. degree from the Indian Institute of Astrophysics, India, in 2001. He is currently a Professor with the Indian Institute of Information Technology Allahabad. His research interests include human gait analysis, human prosthetic, bio-metrics, image processing, graphics and visual computing, graphical projections, robotics and instrumentation, real time simulation, high performance computing (HPC), artificial life simulation and intelligence, and human-computer interaction.